Dorothee Wiegand

Foreman

Machine translation on the PC

To process various texts equally well, the software translating manuals, operating instructions, E-mails or web pages should possess the quality of "pantophagy". For this purpose, the user's skill is invaluable.

The machine translation practices total to decades. While the current language-recognition software often work based on statistics, the major part of the translator applications are based on the rules: they interpret the sentences from the linguistic point of view. The thing is that most difficulties occur when poly-semantic words are to be translated. For example, in German the homonym "Schimmel" may denote as a white horse, so thin greenish coating on stale products; and this is but one of the problems to be adequately translated. Often the differences are much subtler. Depending on the context, the English verb "to take" corresponds as to the German verb "holen" (Atem holen - to take a breath), so "to bring" meaning "to deliver" (I'll take you to the station), to say nothing of such constructions as take away/back/off/on/over, etc.

The polysemy could be tackled with, though, using statistic methods operating on the probability values. For example, for the word "Schimmel" the probability of translations as "horse" could be higher due to immediate proximity to the words "trot" or "gallop". On the other hand, if the word appeared near such words as "cheese" or "sausage", then another meaning would have been determinant.

Nowadays, the adherents of statistic methods and those of methods based on the rules really fight for the "true faith". Some experts are sure, though, that the future machine translation idea would not do without statistic processes [1]. But the statistics is far from the software so far. Instead, it is the user who is to take up the whole of the delicate work.

Our choice of the candidates, among all, was governed by the fact of their ability to influence the translation. Thus, along with the system dictionary, they should have provided the user's dictionaries, where the user could save special terms on his favorite themes. The following five programs that came to our test provide sufficient user participation: @promt Professional from PROMT, Personal Translator 2002 from lingu-atec, Reverso Pro 5.0 from Softissimo, Systran 4.0 from Systran company and T1 5.0 from Langenscheidt. The new version of the earlier Lernout&Hauspie-Produkt Power Translator should have come out as "Logomedia Power Translator", but was not ready before this article was printed.

We tested the versions for English-German translation. Apart from all, @promt provides a version for Russian; there is the French version of Personal Translator, the Langenscheidt software not only for French and Russian, but also for Spanish (newly

issued). The assortment of other languages needs expanding, but there are special applications dealing with exotic languages; though, they lack many useful functions [2].

All the candidates tested offer a translator displaying the source text and its translation behind or beneath. The base case also includes plug-ins for the office programs. As for the unconnected applications, you first have to copy the text into the clipboard before translating it. All programs support Internet Explorer, and some even Netscape Navigator.

The automatic recognition of the source language may help the active users in their work. The synchronization function helps to navigate the bigger documents: when you click of the source text position, the cursor in the translation area goes to the corresponding position in the text of the translation, and vice versa. A function excluding some parts of the text from the translation helps avoid translating, for example, "Musterstadt" in postal addresses as "sample city" or "pattern town". All functions of the programs being tested are given in the comparison table on page 183.

The programs that employ translation memory used to be the most flexible. These are data bases for pairs of sentences such as "Das Ergebnis der Untersuchung ist positiv ausgefallen/The result of the survey was encouraging". Some programs have filled archives the user can supplement, the others are supplied with empty databases. In both cases, by means of thorough corrections and archiving his work the user can constantly improve the translator software.

But he should learn how to use the archive of the sentences. The level of similarity of the source sentence and a variant from the archive is determined by letter or word comparison. Usually, if the source sentence and that from the archive are fully identical,

The latest achievement of @promt is the Plug-in for Acrobat Reader, which translates texts from the documents in the PDF format.
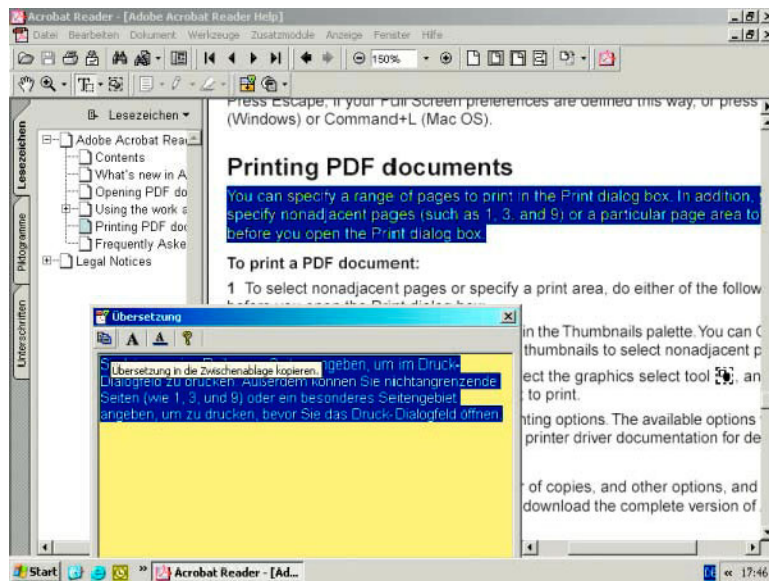
the translation is done via the archive. The user can reduce the level of match down to about 80%. In this case, a sentence from the archive will be translated matching only for 80%. Thus, instead of the "Untersuchung" (survey) the source text could contain "unserer Untersuchung" (of our survey), and the above sentence from the archive would have nevertheless been used.

If the match value is set very liberal, then the sentence "Das Ergebnis … ist negativ ausgefallen" (The result was not encouraging) will be treated as satisfactory, and translated into English as "The result … was encouraging", although the meaning of the sentence in German is diametrically opposite to that in English. The user employing the archive should be aware of such perfidy and view the translated text attentively to avoid "misunderstanding".

Suitable for everyday use

Often the translator software was scientifically tested: with the examination texts for school-leaving certificates and state exams, by educational counsels, native speakers and also on the computer; much attention was paid to confirm that the test result obtained

automatically is equivalent to that obtained by the human tester [3, 4]. Our test (see the table on page 182) is made up not of special documents but of common texts, as we encounter with them in everyday life. All programs were tested on the PC with the 1,2-GHz Athlon processor and 256 Mb RAM.

Almost all tested programs exist in various versions. Versions with the sentence archive included were considered in more detail. The developers' web sites contain tables with meticulous listing of the differences among the versions. They compile the software reference list for this article, with the web-sites where you can test the translator skill of the software in the online mode.

@promt Professional

So far the Russian manufacturer PROMT cooperated with the Softissimo company from Paris. Their Reverso product also works with the translator system developed in St. Petersburg. The market was divided: the French company was responsible for development and sale of the English-German and French-German translators. Now the ways of the two companies seem to be separated and, moreover, they became competitors at least

regarding the English-German language pair. The professional version includes the sentence archive; in the Expert version the developer additionally installed the terminology manager as well as TRADOS translator database interface.

The main application consists of the editor that can alternatively represent the source text and the target text combined with the lists of unknown words in various representations and offers flexible printing options. For the purpose of additional processing another useful option is the ability to save the source text and the target text in a table, where the source and translated paragraphs are positioned side by side.

The translation settings in @promt enable to connect to the packages later assigned to certain themes. For example, the "manual" type of text as a list of actions can be assigned, first, the formal "You", second, the translation of the English imperative as polite infinitive ("Please, turn the switch in the start position") and, third, the user compiled technical archive of the sentences.

It is necessary to select a lot of functions in the menus "Theme", "AM" ("Associated Memory") and "Translation", and there are cases when neither comprehensive man-

Translator software is powerless to puns or irony. The electronic edition of the Langenscheidt pocket dictionary integrated into T1 can provide competent help.
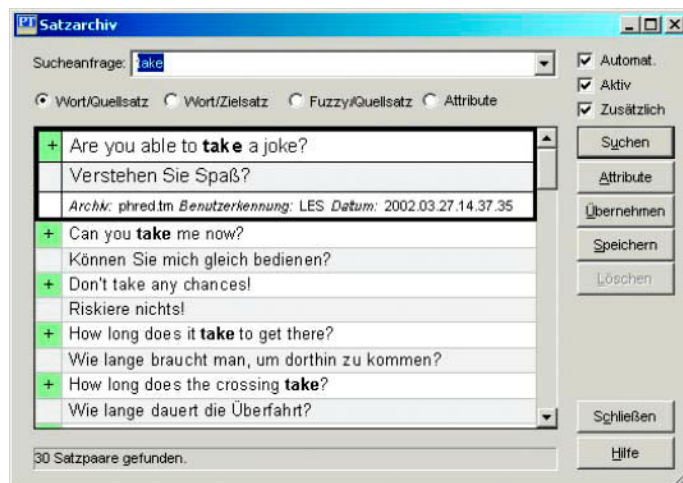
ual, nor online Help can be useful. When the main application is mastered, the following sections of the program are worth considering: the software comes with the Dictionary Editor, Archive Manager, Backup facility and the mini-translator "PromtX".

When tested, @promt demonstrated the average quality of the translation as English-German, so German-English. As for the integration with Microsoft Word, @promt proved to be the only fully suitable program among those tested: the target text contained correctly translated and formatted references with related text, as well as the page numbering including the date.

@promt Professional provides the plug-in for the Adobe family software enabling PDF format texts translation. The availability of the sentence archives is clear advantage of the software, even though now it is delivered with empty database. Practical accessories like Backup facility speak for many useful ideas of the developers. To facilitate maintenance, in the future versions the developers are to practically combine these accessories, as well as modify and expand the context Help.

Langenscheidt T1 Professional 5.0

For the English–German languages the Langenscheidt software is offered in three variants: Standard, Standard Plus and the professional edition tested here — the one with the *translation memory*. T1 stands for "Translation archives". The software is delivered with two archives: "Business" with 5200 sentence pairs and "Idiom" with almost 71 000 entries; the user can create or import additional entries on his own. Among the business entries there are patterns for business letters. Some of them are very specific, e.g., „Als Hersteller von vollautomatischen Waschmaschinen haben wir einen groЯen Bedarf an elektrischen Motoren von 0,1 bis 0,5 PS" together with the English match: "Being the manufacturer of the fully automatic laundry washers we are in constant need of electric motors 0.1 to 0.5 h/p". Mostly, however, the suitable general formulations are considered in the form they are used in modern correspondence.

The software failed to translate correctly the phrase "Intensiv-Sprachkurs" (intensive language course), whereas there were no problems with the correct wording "Intensivsprachkurs", according to Duden. The overall quality of translation made by T1 during the test was below the average, which may explained by scant stock of abbreviations. With the specialized text in English the software demonstrated better result.

Using the configuration dialog box T1 can be activated and deactivated in certain components of Microsoft Office. If the option is checked, then the additional buttons and the new "T1" menu appears in Word, Excel, Powerpoint and Outlook. There is also a plug-in for Access. Those who use the T1 services from one of the Office applications may later discover some remnants on the hard disk. In the course of work with Word the user is not asked anymore, whether he would like to save his files: the plug-in blocks this message and saves up to 3 files with the target texts, as well as with various additional information in the folder of the source text. If a mail is translated in Outlook, then the copy of the source message is saved on the IMAP server. The transmission of a HTML file via the main application left the target text as the RTF file, unlike in the Internet Explorer, where it was transmitted very accurately.

The electronic version of the pocket English–German Langenscheidt dictionary, where search is performed using "T1-Lookup", also refers to the software. The small search program can also be connected to the other electronic dictionaries from Langenscheidt. The users having one or more of such special dictionaries installed can perform search in them from T1. The German manual

including the detailed list of keywords in the end is introduced only in the PDF format, and there is no online Help there.

Personal Translator 2002 Office Plus

The translator software from Linguatec is also introduced in the "Home" variant for the beginners; the more advanced versions for one user are called "Office" and "Office Plus".

If the Microsoft Word interface and the availability of the sentence archives matter, then the Plus version tested here is the right choice to recommend. When just installed it contains 1400 general phrases for German-English translation; for English-German translation there are 1138 common phrases and sentences provided.

In our test the Personal Translator produced relatively better results compared to the others. The quality of the German-English translation of the test files granted this software the 1st place; as for English-German, the translation was of average quality. In E-G translation the program left many abbreviations unprocessed, despite that both messages, and linguistic texts were translated very well. In our test file in the DOC format the graphics, tables, page numbering and references were retained, though the text of the reference was remained in German. It took quite a lot of time for the program to process the test file. Large size files are worth diving into smaller pieces, and the "Background translation" option offered by the program can also be useful.

The user can supplement the sentence archives of the Personal Translator as he likes. To find an optimal personal setting the user has to experiment with the offered options. For example, a 60% match in the test turned out to be a rather liberal match value: the translation of the test sentence "Grass is always

greener on the other side" with the use of the installed archive became: "Are there still places on an earlier flight?" Obviously, the comparison algorithm found linguistic similarity between absolutely different sentences.

The HTML files are translated only via batch processing. Besides, usually reliable automatic recognition of the translation is not working here. Despite that the program recognized correctly the meaning of the "liesmich.html" test file: "The language of the page is possibly German", its second part was interpreted really cute. If the translation direction is set manually, then HTML works well. However, no reference links are retained within the HTML file.

Due to informative manual, perfect online Help, integrated editor recognizing outdents and millimeter-exact leading Personal Translator may be treated as really powerful translating tool. Sometimes working with Windows/Office XP the program buzzed. Hopefully, the next version of the good translator system will feature more sophisticated interfaces.

Reverso Pro 5.0

After installation and mandatory restart the Reverso user finds three new icons on the desktop: PROMT-L, WebView and QTrans. QTrans is intended for quick translation of the unformatted texts. In the split window of the browser WebView the user can see the original page and its translated variant. The translation quality varies according to complexity of the template. Nevertheless, WebView does not ruin the page layout and works rather quickly. Those preferring the original Internet Explorer will find the additional translator buttons there.

The control panel of the main application PROMT-L looks very much like the PROMT software. In the test the Reverso program was the winner in translation into German, but the last one in trans-

The sentence archives of the Personal Translator from Linguatec recognizes 50 idioms with the verb "take".
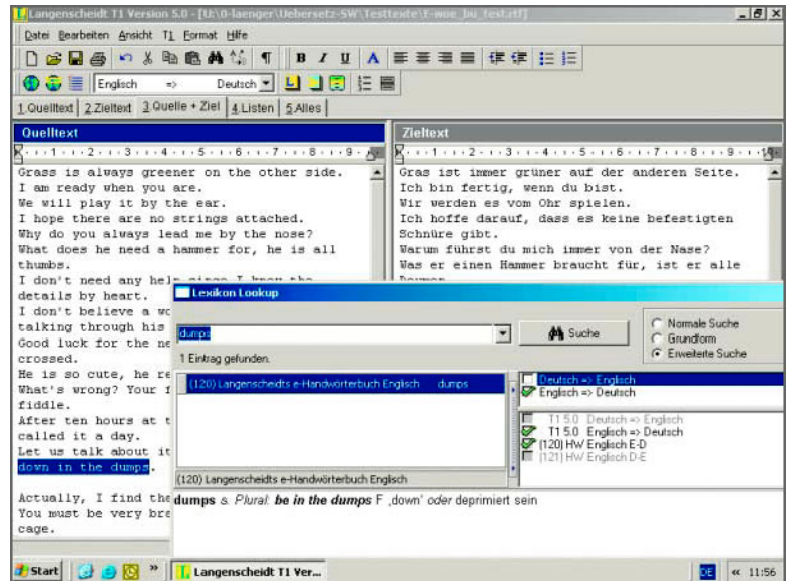
lation from German. The functional range of the main application is rather scanty, but the existing functions are very rationally positioned in the menu, so even the beginners find it easy to master the program. Here is also introduced the useful option of saving the source and target texts together in one document.

When the test file was opened in the DOC format (created in Word 2002), its initial format was almost completely lost. The tables, graphics, references and page numbering were lost. Translation of same file in the RTF format took rather long time, and there were also mistakes,

for the references and page numbering were lost. The Excel-Plug-in, on the contrary, worked perfectly well.

The dictionary is unaware of the new rules of German spelling: it does not understand the word "essenziel" or writes the verb "nummeriert" with one "m". By the time this article was approved for publishing it was found out that Reverso was implemented only as the combined English/French edition.
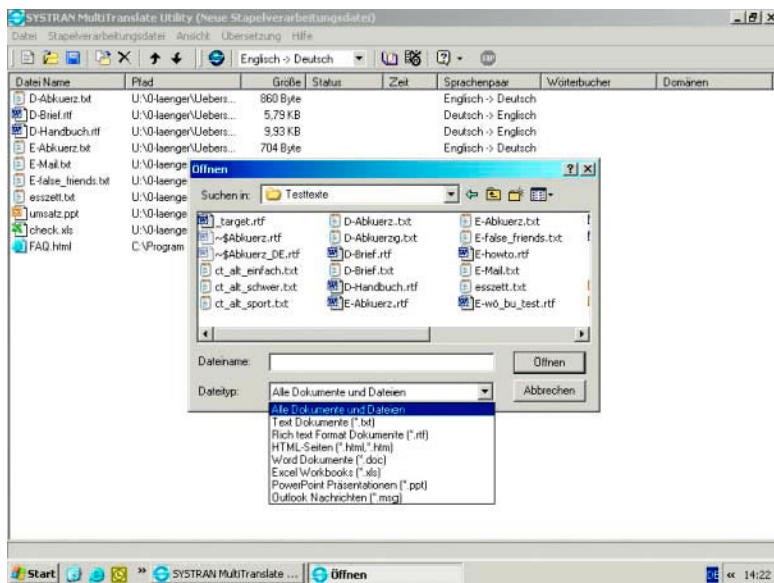
Systran 4.0

The translator system from Systran is the old-timer among the tested programs. The be-

ginning of this software initially developed for high-performance computers goes back for more than 50 years. It has been used for decades in the European Union and is hidden behind many links in the Internet marked, for example, as "Click here for translation" and for widely used languages in the translation function of Google.

At first sight, the menu bar seems attractively self-explanatory, but after the word "ascetic" appeared more appropriate. For example, the "Format" menu contains only font selection and paragraph alignment: center, left, right. It is more convenient to select
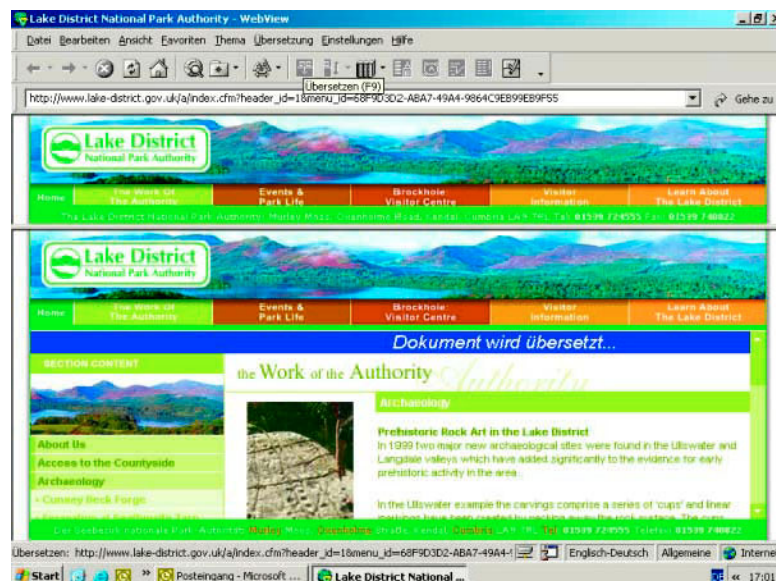
For bilingual search Reverso has "WebView" — customized version of Internet Explorer taking over the existing settings and Favorites.

Systran Premium contains "MultiTranslate Utility" working on the background. Besides, it is possible to set the translation direction separately for each file.

both on the icon bar. The program Help is also scanty; the user cannot search by keywords and obtains no assistance when pressing F1.

The translation performance of Systran turned out to be good in the test, and the software occupied the second place as in E-G, so in G-E translations. The program translated "Herr Dr. Müller" as "Dr. Mueller", not as "Mister Dr. miller", like the other test participants. When translating the letter from German the sense was mostly retained, but the program encountered with certain difficulties in translating composite words.

The standard version of Systran failed with the 60 pages of the RTF file; the latter was loaded, but during translation a system failure of Systran occurred. Fortunately, the Premium variant has the so-called MultiTranslate Utility the user can employ to make up a list of any files to be later translated as a batch. Apart from TXT, RTF and DOC, also the HTML, PPT, MSG and XLS formats are allowed. In this case the translation of both files was not only quick, but also problem-free, either.

The speed and quality of the translation are self-ever, the software may be considered obsolete. The new German spelling has not been introduced into the dictionary yet, and the control panel does not meet the contemporary ergonomic requirements.

Results

None of the tested programs can supplement lack of English language knowledge. Only the simple sentences without linguistic traps were translated absolutely correctly by all applications. The problems begin where slightly "rusty" school English fails. Of course all tested programs did translate our short letter to a foreign language school, but not without inadvertent comic effect and a few blunders. One would hardly wish to make business abroad in such a style.

Those who translate texts occasionally can do it better manually; indeed, decent results need such a profound preparatory work and "polishing" that use of the computer saves no time. All this is arguable, though, if the user is interested only in general sense of a document. All these program tests allowed to coarsely find out if English inquiry of the product was worth developing, and whether there would be market demand for them.

The translator systems may be really interesting for the professional who often translate commonly used texts. First of all, it is the sentence archive that helps most when user guides or manuals are to be translated, where standard phrases are not uncommon. Such are offered by the full versions of @promt, Langenscheidt T1 and Linguatec Personal Translator. While @promt starts from "blank page", the archives of T1 and Personal Translator already supported with sentence pairs. Those who have the electronic dictionaries from Langenscheidt, may want to consider T1 in more detail. On the other hand, @promt is easier to master, but it is intuitive Personal Translator, which is the easiest program.

If one can renounce sentence archives, then Reverso is also in the range along with the 3 variants of the mentioned programs. With this software a beginner can achieve good results. The translation quality in the test was good, but Reverso abilities were limited as of affecting the translation process. Systran seems to be more suitable for patented networks, than for the PC.

For the sake of feasibility, all developers provided plug-ins as for the Microsoft Office components, so for those of the Internet Explorer. Here "less" would often mean "more". Instead of providing more or less operable links with maximal number of the programs, the developers should have first optimized the work with Word, where only @promt demonstrated perfect performance.

When translating web-sites all participants of the test compete a with the translation service of Google. From the linguistic point of view the result is often just satisfactory or grotesque, but it is easy to use, and the page layout is not ruined. Only Reverso and @promt demonstrated the best results in the test. (dwi)

Literature

[1] Hermann Ney, Maschinelle Sprachverarbeitung, Der statistische Ansatz bei Spracherkennung und Sprachübersetzung, Informatik Spektrum 26(2), 2003, S. 94-102
[2] www.alamax.de/service/ueberblick4.php3
[3] www.nist.gov/speech/tests/mt/doc/ngram-study.pdf
[4] www.cs.mu.oz.au/acl/P/P02/P02-1040.pdf

Soft-Link 0 321 174 **c't**

Time needed for translation of 60 pages (sec.)

	< better	
@promt 6.0		93
Personal Translator 2002		135
Reverso 5.0		76
Systran 4.0		14
T1 5.0		58

Looking for all-round talent

When translating from English into German, the test candidates should have "fight" the abbreviations, data, and adequacy using the set of sentences, and then prepare the text more accurately for the subtleties of the translation. Finally, there was a task of translating the e-mail message with description of a fictitious software "Wondersoft XXL".

In the opposite direction (English-German) the tested programs were offered the sentences with widely used abbreviations and dates, as well as a letter and several pages of the manual. The letter to the English linguistic school deals with the proposals of the intensive courses. The manual text begins with a table of contents, and in the text part it is divided into the numbered pieces. Apart from possibly good translation for the tested program it was important to convey the numeric system, as well as insertions and tabs.

Finally, we checked the retention of various formats and fonts of the source text, as well as the inserted tables, diagrams, references and page numbering. A few random tests helped us to check the knowledge of the old and new German spelling, as well as differences of British and American word usage. The tested programs also should have translated a 60-pages RTF file on time. If a program had the plug-in for text processing, table calculation or web-browser, these features were also tested.

When the translation quality was estimated, each free-and-easy statement clear in the target text brought the participant 1 point; the translations of the best quality gained additional points. The following table demonstrates the ability of the contemporary translator systems and the need of their improvement.

Source and target texts

	English E-mail	English subject text	Abbreviations, Units	Letter in German
Original	Dear Ulrike, Just a quick email to keep you informed about our latest products and developments.	The increased usage of English loan words in German may not be as much of a boon to the English speaker as it may appear.	Im August wurden mehrfach 42° im Schatten gemessen. / For details see fig. 6.	Sehr geehrte Damen und Herren! Ich interessiere mich für einen Intensiv-Sprachkurs in Großbritannien. In der Schule hatte ich zwar sechs Jahre EnglischUnterricht.
@promt	Liebe Ulrike, Gerade ein schnelles{promptes} E-Mail, um Sie informiert über unsere spätesten {letzten} Produkte und Entwicklungen zu behalten.	Der vergrößerte Gebrauch von englischen Kreditwörtern {Kreditworten} kann nicht auf Deutsch soviel eines Segens dem englischen Sprecher {Lautsprecher} sein, wie es erscheinen kann.	In August 42 were measured repeatedly ° in the shade. / Für Details sieh Feige 6.	Dear ladies and gentlemen! I am interested in an intensive linguistic course in Great Britain. Indeed in the school I had six years of English lesson.
Personal Translator	Liebe Ulrike, Nur eine schnelle E-Mail, um Sie zu behalten, informierte über unsere letzten Produkte und Entwicklungen.	Es kann sein, dass der gesteigerte Gebrauch von englischen Lehnwörtern auf Deutsch nicht ebenso viel von einer Gnade zum englischen Sprecher ist, wie er erscheinen kann.	42° were repeatedly measured in the shade in August. / Für Details sehen Sie Fig. 6.	Dear Sir or Madam! I am interested in an intensive language course in Great Britain. In school, although I had six years English lessons.
Reverso	Liebe Ulrike, Gerade ein schnelles EMail, um Sie informiert über unsere spätesten{*letzten*} Produkte und Entwicklungen zu behalten.	Der vergrößerte Gebrauch von englischen Darlehen-Wörtern kann nicht auf deutsch soviel einer Gabe dem englischen Sprecher sein, wie es erscheinen kann.	In August 42 were measured repeatedly ° in the shade. / Für Details sieh Abb. 6.	Dear ladies and gentlemen! I am interested in an intense language course in Great Britain. Certainly, in the school I had six-year English lesson.
Systran	Liebe Ulrike, Gerade ein schnelles email, zum Sie zu halten informiert über unsere neuesten Produkte und Entwicklungen.	Der erhöhte Verbrauch der englischen Darlehenwörter auf Deutsch kann möglicherweise nicht so viel einer Gabe zum englischen Lautsprecher sein, wie er erscheinen kann.	In August 42° in the shade were measured several times. / Für Details fig. 6 sehen.	Ladies and Gentlemen! I am interested in an intensive language course in Great Britain. In the school I had six years English instruction.

T1	Teures <U[Ulrike]>,Dass ein schnelles <U[email]> dich gerade über unsere spätesten Erzeugnisse und Entwicklungen informiert hält.	Der vergrößerte Gebrauch von englischen Darlehenswörtern in Deutsch<A[Deutsch|Deutschem]>darf nicht soviel eines Segens zum englischen Sprecher sein, wie es erscheinen kann.	In August 42 ° were measured repeatedly in the shadow. / Da Details Abb. 6 sehen.	Dear Sir or Madam, I am interested in an intense-Sprachkurs in Great Britain. In the school I had indeed six years of English-instruction.

Translator systems – check list

	@promt	Personal Translator 2002	Reverso 5.0	Systran 4.0	Langenscheidt T1 5.0
Developer	PROMT	linguatec GmbH	Softissimo	Systran	Langenscheidt
Sale	abitz.com/Jourist	linguatec /торговля	Softline AG	Heisoft	Langenscheidt / Handel
URL	www.abitz.com www.jourist.de	www.personal-translator.de	www.softline.de	www.heisoft.de	www.langen scheidt.de
Telephone	0 30/68 47 290 0 40/21 09 82 90	0 89/89 66 640	07 81/92 93 222	02 21/94 26 73 33	0 18 05/23 23 57
System requirements	Windows 95 to XP	Windows 95 to XP	Windows 98 to XP	Windows 98 to XP	Windows 98 to XP
Equipment					
User's manual/Online-Help/context help F1	dt./dt./v	dt./dt./v	dt./engl./v	engl./dt./–	dt./dt./–
Specialized integrated dictionaries/available on request	–/–	–/v	–/v	–/v1	–/v
Sentence archives	v	v	–	–	v
OCR software	–	v	–	–	v
Speech output	(v) [2]	v	(v) [2]	–	–
Integration in other applications					
Word	v	v	v	v	v
Excel/PowerPoint/Access	v/v/–	–/–/–	v/v/–	v/v/–	v/v/v
Internet Explorer/Netscape Navigator	v/–	v/v	v/–	v/v	v/–
Outlook	v	–	v	v	v
Installation options					
D->G: pronouns alternatively m/w / automatic recognition	–/–	v/v	–/–	v/–	–/–
Alternative translation of words in capitals	–	–	v	v	v
Infinitive translation alternatively as an imperative	v	v	–	v	v
They -> You/They or You -> Them (Her,It) / You alternatively	v	v	v	–	v
Alternative division of composite words	v	v	–	–	–
Alternative paragraph marks as end of sentence interpretation	–	v	–	–	–
Old or new spelling, alternatively	–	v	–	–	v
Distinction of British/American English	v	v	v	–	v
Translation					
Automatic recognition of source language	v	v	v	–	–
Synchronize function	word-wise	phrase by phrase	in paragraphs	phrase by phrase	–
Expulsion of text passages from translation	v	(v) [3]	v	v	–
Time limit per sentence	–	v	–	–	–
Possibility of translation process cancelation	v	v	v	–	v
General program maintenance					
Views: Source/Target/both/others	v/v/v/v	–/–/v/–	–/–/v/v	–/–/v/–	v/v/v/v
Print options: Target/Source/both/unrecognized words	v/v/v/v	v/v/–/–	v/v/v/v	–/–/–/–	v/v/v/v
Send target text to Mailclient	v	v	v	–	–
Maximum number of undos	1	1	1	> 50	> 50
Possible import of term lists	–	v	–	–	v
Estimation					
Equipment	++	++	±	-	+
Integration with other applications	++	±	+	+	+
Installation options	+	++	±	±	+
Translation	+	+	±	±	±
Program operation	+	+	+	±	+

Prices

	Office 99 €	home 49 €			Stand. 159 €
	Profess. 299 €	Office 99 €	Combination issue	Standard 459 €	Stand. Plus 209 €
	Expert 599 €	Office Plus 249 €	Eng./French. 398 €	Professional 1159 €	Profess. 299 €

[1] Premium version contains 21 specialized dictionaries   [2] Uses Windows speech output   [3] only Word Plug-in

++ very good   **+** good   **±** satisfactory   **-** bad   **--** very bad   **V** existing   **–** not existing   k.A. no Information